

EMEDIA		EMEDIA LIVE!	EMEDIA MAGAZINE	DVD PRO	ONLINE INC.
CURRENT ISSUE	SUBSCRIPTIONS	EDITOR'S CHOICE	EDITORIAL CALENDAR	INFOLINK	CONTACT


Get The Latest **Digital Audio** News

[Browse by Subject](#)

information insider

Why Bother with XML? Here's Why.

Robert J. Boeri
Martin Hensel

EMedia, January 2000
[Copyright © Online Inc.](#)

If you've been unconvinced by the XML hype--comparisons to ASCII as the new text touchstone and the like--you're not alone. In general, it's good to challenge any new standards (or family of standards), especially when the adoption encouragement is coming from vendors wanting to sell you the "next big thing." Remember: the World Wide Web Consortium (W3C) is a vendor consortium, not an international-standards body like ISO. In fact, the effect of the W3C's XML initiatives has apparently surprised even the W3C: XML for non-document Internet applications (such as ebusiness) has not waited for some of the newer standards to emerge, since data interchange doesn't require some of them.

If your business is electronic publishing--Web or optical--there are a good many arguments, for the time being, to stay with familiar tools and document formats. But the time is now to start exploring XML and its various related standards, and to realize that Web browsers--increasingly XML-aware--are rapidly becoming the interface of choice for applications and data on both optical media and the Web.

word processing to web-CD delivery?

Several electronic book-publishing system vendors can accept word-processing files instead of (or in addition to) XML or SGML. Examples include Hynet's Directive and Enigma's Insight and Dynatext (Enigma and Inso recently agreed that Enigma would acquire Inso's DynaText and DynaWeb product lines). Such products typically accept Microsoft Word and Corel WordPerfect; some accept FrameMaker and even Acrobat PDF files.

But read the fine print in these vendors' electronic-book delivery offerings. Want automated links? Zoned full-text searching? You can get these. However, all systems are limited by the quality of source documents; "garbage in, garbage out" applies big-time. Word processors are designed to produce excellent-looking documents, but they are blind to structural defects introduced by users. For example, in any word processor you can produce a table-like form in several ways: use a table command, create columns via tabs, and simulate columns with multiple spaces separating them. Each technique may produce good-looking printed documents, but each will not produce identically functional electronic books.

Suppose you've recognized the need for discipline and developed uniform-style templates for your organization. This technique can make authoring documents easier (e.g., authors could use a copyright style instead of remembering where the "copyright" pie font is), and styles can add a level of visual consistency. If styles are applied consistently, electronic books can use them much as they would XML tags. Unfortunately, word processors cannot assure that anyone uses style libraries correctly. Corporate style guides may dictate no level-three heading immediately following a level-one heading, but word processor-style libraries can't enforce that. Since electronic book-publishing systems also use and depend on style elements, if those elements are misused, the results will suffer, too.

increasing your vocabulary

XML doesn't require document models provided you follow some basic rules. And there can be problems building document models-- indeed, getting various industry groups to create them and then agree to abide by them is difficult. However, industry-standard vocabularies make applications and documents more productive and rules-based, and standard vocabularies are being built today. Microsoft's BizTalk (<http://www.biztalk.org>) is encouraging industry groups to develop standards for business-to-business and data interchange, but it is also likely that when those schemas are agreed upon, they will be useful in XML documents. ACORD, a standards-setting consortium for the insurance industry, is creating insurance models. XML Open is also coordinating industry initiatives (see http://www.xml.org/xmlorg_catalog.htm for a list of participants). Industry-specific documents published on the Web or optical media will have to conform to these models.

Do you want to build a model once and express it in many different formats, present and future? A new \$100 tool, "XML Authority," lets you build a model once and save it as: A DTD; XML Schema Definition Language (the emerging W3C specification); DCD, the Document Content Description for XML; XML-Data; the Document Definition Markup Language (DDML); and SOX, the Schema for Object-oriented XML.

And getting started can be easy. You can use many existing structured sources of structured information to jump-start creation of an XML model. XML Authority, for example, accepts a variety of sources to jump-start your model development. Inputs include: existing XML-marked documents; SGML DTDs; column headers from delimited text files; Java classes (with the class becoming an XML element and its properties becoming elements or attributes); COM objects; ODBC table schemas; and LDAP directory structures. Never mind if some of these aren't familiar to you. The important point is that you're probably working with structure already, and tools like XML Authority can help you get into the world of XML rapidly. For a try-out copy, visit <http://www.extensibility.com>.

Browsers are becoming XML-aware. How about you?

Robert J. Boeri (bboeri@world.std.com) and *Martin Hensel* (mhensel@texterity.com) are co-columnists for *Information Insider*. Boeri is an Information Systems Publishing consultant at a Boston-area insurance company. Hensel is president of *Texterity, Inc.*, a Newton, Massachusetts-based consulting firm that builds SGML-based editorial and production systems for publishers, corporations, ecommerce services, and type-setters.

Comments? Email us at letters@onlineinc.com.

[\[EMedia Home\]](#) [\[Subscriptions\]](#)

Copyright © 2000, [Online Inc.](#) All rights reserved.
info@onlineinc.com